

## Solución a los ejercicios de autocomprobación

1) La clasificación de las variables registradas es la siguiente:

- i) Número de días lluviosos en un mes: cuantitativa discreta, medida en una escala de intervalo/razón.
- ii) Estación del año: primavera, verano, otoño, invierno: variable cualitativa medida en una escala nominal.
- iii) Temperatura máxima diaria en ° Centígrados: variable cuantitativa continua medida en una escala de intervalo/razón.
- iv) Número de rayos caídos en un mes: cuantitativa discreta medida en una escala de intervalo/razón.

2) Se presentan las cantidades que faltaban en la tabla:

Valores	$n_i$	$N_i$	$f_i$	$F_i$
10	2	2	0,05	0,05
20	10	12	0,24	0,29
30	8	20	0,19	0,48
40	6	26	0,15	0,63
50	8	34	0,19	0,82
60	4	38	0,10	0,92
70	3	41	0,08	1
<b>TOTALES</b>	<b>41</b>		<b>1</b>	

Tabla 14: Tabla solución del ejercicio 2

3) Distribuciones marginales para X e Y.

X	$n_i$	Y	$n_i$
[0,2)	38	1	42
[2,4)	19	4	24
[4,6)	57	6	30
TOTAL	114	8	18
		TOTAL	114

Tabla 15: Tabla solución del ejercicio 3 (marginales)

Distribuciones condicionadas:

<b>X Y=6</b>	<b>n<sub>i</sub></b>	<b>Y X=[2,4)</b>	<b>n<sub>i</sub></b>
<b>[0,2)</b>	10	<b>1</b>	7
<b>[2,4)</b>	5	<b>4</b>	4
<b>[4,6)</b>	15	<b>6</b>	5
<b>TOTAL</b>	30	<b>8</b>	3
		<b>TOTAL</b>	19

Tabla 16: Tabla solución del ejercicio 3 (condicionadas)

Se estudia ahora la independencia de las variables X e Y. Es fácil comprobar que en toda la tabla de frecuencias bivariantes se verifica:

$$n_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}$$

<b>X / Y</b>	<b>1</b>	<b>4</b>	<b>6</b>	<b>8</b>	<b>Marginal X</b>
<b>[0,2)</b>	14	8	10	6	38
<b>[2,4)</b>	7	4	5	3	19
<b>[4,6)</b>	21	12	15	9	57
<b>Marginal Y</b>	42	24	30	18	114

Tabla 17: Tabla completa para ver independencia

Por ejemplo,

$$14 = \frac{38 \times 42}{114}, \quad 8 = \frac{38 \times 24}{114}, \quad \dots, \quad 7 = \frac{19 \times 42}{114}, \quad \dots, \quad 21 = \frac{57 \times 42}{114}, \quad \dots, \quad 9 = \frac{57 \times 18}{114}$$

Por tanto, las variables X e Y son independientes.

- 4) Se estudia la variable  $X = \text{Condición de fumar}$  y la variable  $Y = \text{Sexo}$  en una muestra de 52 estudiantes:
- Fuman 19 de 52 estudiantes, un 36%.
  - No fuman 17 chicas de un total de 22, un 77%.
  - Son chicos y fumadores 14 estudiantes de 52, un 27%. El resto de los estudiantes son chicos no fumadores (un 30%), chicas fumadoras (un 10%) y chicas no fumadoras (un 33%).
  - Distribución marginal de frecuencias relativas de la variable  $Y = \text{Sexo}$ .

Sexo	$f_i$
Chicas	0,42
Chicos	0,58
<b>TOTAL</b>	<b>1</b>

Tabla 18: Tabla solución del ejercicio 4d)

- e) Distribución marginal de frecuencias absolutas de la variable  $X = \text{Condición de fumar}$ .

Condición de fumador	$n_i$
Fuma	19
No Fuma	33
<b>TOTAL</b>	<b>52</b>

Tabla 19: Tabla solución del ejercicio 4e)

- f) Distribución de frecuencias relativas de la variable  $Y = \text{Sexo}$  condicionada a que son alumnos fumadores.

Sexo Fumadores	$f_i$
Chicas	0,26
Chicos	0,74
<b>TOTAL</b>	<b>1</b>

Tabla 20: Tabla solución del ejercicio 4f)

5) Una posibilidad para representar el diagrama de tallo-hojas es redondear los datos perdiendo la cifra de las centésimas: 4,5; 5,6; 4,1; 5,2; etc. La hoja será el segundo dígito y el tronco el primero. Esta forma de redondear los datos daría lugar a un diagrama de solamente dos tallos:

```

4 | 1 1 2 2 2 3 3 3 3 4 4 4 5 5 5 6 6 6 7 8
5 | 0 2 5 6 6 7

```

Este diagrama es poco informativo. Una técnica para extender este tipo de diagramas es dividir cada fila en dos o cinco partes iguales. Si se eligen cinco partes iguales se utilizan las siguientes etiquetas:

- ✓ \* para las hojas que llevan las cifras el 0 y 1
- ✓ T para las hojas que llevan las cifras 2 y 3 (Two, Three)

- ✓ F para las hojas que llevan las cifras 4 y 5 (Four, Five)
- ✓ S para las hojas que llevan las cifras 6 y 7 (Six, Seven)
- ✓ ● para las hojas que llevan las cifras 8 y 9

El diagrama de tallo-hojas quedaría de la siguiente forma:

4 *	1 1
T	2 2 2 3 3 3 3
F	4 4 4 5 5 5
S	6 6 6 7
●	8
5 *	0
T	2
F	5
S	6 6 7

en el que se aprecia con más claridad el perfil de la distribución.